Master's Thesis

A survey of the optimal time series modelling for newbuilding ship prices

Jae-Min Jeon

Graduate School of Technology and Innovation Management

UNIST

2020

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

A survey of the optimal time series modelling for newbuilding ship prices

Jae-Min Jeon

Graduate School of Technology and Innovation Management

UNIST

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

A survey of the optimal time series modelling for newbuilding ship prices

A thesis/dissertation

submitted to the Graduate School of Technology and Innovation Management, UNIST

in partial fulfillment of the

requirements for the degree of

Master in Technology and Innovation Management

Jae-Min Jeon

12/09/2019

Approved by

Han Woo

Advisor

Han-Gyun Woo

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

A survey of the optimal time series modelling for newbuilding ship prices

Jae-Min Jeon

This certifies that the thesis/dissertation of Jae-Min Jeon is approved.

12/09/2019

signature

Advisor: Han-Gyun Woo

signature

Dae-Jin Kim

signature

Kee-Yeun Lee

Abstract

The shipping industry can be divided into four markets : the freight market, the newbuilding market, the second-hand, and the demolition market. The newbuilding market is especially responsible for connecting two pillars of maritime business—the shipping and shipbuilding industries—with newbuilding ship prices. The newbuilding ship prices are the sum of measured values of vessels to be constructed at the time of contracting with shipowners, and newbuilding ship price market forecasting would be a criterion of strategic decision-making in shipyards. Therefore, a reasonable estimation of the newbuilding ship price can be a driver for growth in shipyard management.

Previous studies on the determining factors for newbuilding ship prices are rare, and some of the work is old and requires reinvestigated. Also, since the newbuilding market is volatile, time-series forecasting methodologies that assume linearity have limitations in terms of utilization. To propose an optimal newbuilding ship price estimation model, we built and compared Vector Error Correction Model (VECM) Long Short-Term Memory (LSTM) with hyper-parameter optimization. Through a literature review, we selected economic variables, including second-hand ship prices, freight rates, and interest rates from January 1986 to June 2019, and verified their influence on newbuilding ship prices. For the validation and evaluation of the time-series models, we conducted a sliding window test to achieve prediction robustness. As a result, we empirically confirmed the superiority of LSTM based on neural network that revealed better performance in rapidly changing periods.

Additionally, we applied a Savitzky–Golay filter that eliminates noises from time-series variables and combined it with the forecasting models, and the experimental results indicate that models that are integrated with denoising filters exhibit better performance than single models. Based on empirical tests in this study, we propose a time-series forecasting model combining the Savitzky–Golay filter and LSTM in the newbuilding ship price market.

Keywords : LSTM, Savitzky–Golay filter, Newbuilding Price Index, VECM

Contents

List of Figures

List of Tables

Abbreviations

Adam	Adaptive with moment
ADF	Augmented Dickey–Fuller
AIC	Akaike Information Criterion
ANN	Artificial Neural Networks
ARCH	Autoregressive Conditional Heteroscedasticity
ARMA	Auto Regressive Moving Average
BDI	Baltic Dry Index
ECT	Error Correction Term
EMH	Efficient-Market Hypothesis
GARCH	Generalized Auto Regressive Conditional Heteroskedasticity
HQC	Hannah-Quinn Information Criterion
KOSPI200	Korea Composite Stock Price Index 200
LIBOR	London Inter-Bank Offered Rates
LIBOR3	US Dollar 3-month LIBOR
LSTM	Long Short-Term Memory
MAPE	Mean Absolute Percentage Error
MLE	Maximum Likelihood Estimation
NPI	Newbuilding Price Index
PP	Phillips–Perron
RE	Rational Expectations
ReLU	Rectified Linear Unit
RMSE	Root Mean Square Error
RMSProp	Root Mean-Squared Prop
RNN	Recurrent Neural Networks
SIC	Schwarz Information Criterion
SGD	Stochastic Gradient Descent
SG–LSTM	Savitzky–Golay LSTM
SG–VECM	Savitzky–Golay VECM
SPI	Second-hand Price Index
TAR	Threshold Auto Regressive
VAR	Vector Auto Regressive
VECM	Vector Error Correction Model
VLCC	Very Large Crude/oil Carriers

1. Introduction

The shipbuilding industry has been in a long recession since its peak in 2008. Consequently, major shipbuilders reduced labor force and sold facilities. After years of hardship, the shipbuilding industry began to recover in 2018 due to the increase of eco-friendly fuel freight rates and the expectation of shipowners to replace old ships due to environmental regulations.

Shipbuilding prices, which have a significant impact on the shipbuilding market, have increased steadily since the early 2000s when the shipping boom began, and Clarkson's Newbuilding Price Index (NPI) recorded an all-time high of 191.51 as of August 2008, and the Second-Hand Price Index (SPI) rose to 310.90 in July 2008. However, since the Lehman Brothers' global financial crisis in 2008, the ship prices began to drop sharply. As of December 2009, NPI and SPI recorded 137.84 and 126.59, and dropped rapidly to 28% and 59% in just 1 year, respectively.

Even though the shipbuilding industry is commonly perceived as an industry in which the economic contraction and recovery cycles are repeated (Stopford (2009)), uncertainties still exist in the business environment such as higher interest rates and deepening protectionism. Therefore, it is necessary to predict the value of the upcoming ship market and to use it appropriately for future strategic decisions such as determining related policies, and constructing and investing shipyard infrastructure.

This study starts with the following questions for the necessity of forecasting newbuilding ship price market and practical prediction modelling. Firstly, in previous studies, second-hand ship prices, interest rates, and freight rates are shown to affect newbuilding ship prices. Are these determinants still valid. Secondly, which approach has better prediction performance in newbuilding ship price market between Long Short-Term Memory (LSTM) which is based on neural network, and Vector Error Correction Model (VECM) which represents econometric time-series analysis with linear combinations. Lastly, can the prediction performance of newbuilding ship prices be further improved, if the denoising filter used to eliminate noise in the signal processing is applied to economic data related to ship prices.

To answer these questions, we conducted a unit-root test to examine the characteristics of economic variables related to ship prices and verified the causalities between time-series variables with Granger's (1969) approach. Additionally, we built a VECM, which adds an error correction term to the Vector Auto Regressive (VAR) model when cointegration exists. The VECM compensates for the disadvantage of losing information on long-term equilibrium between variables due to the differentiation in multivariate VAR analysis.

Traditional time-series analysis approaches are not free from statistical assumptions such as mutual independence and linearity between variables. In the neural network field, meanwhile, time-series forecasting methodologies have emerged to cope with these shortcomings. Neural network can be a powerful tool with self-learning mechanisms to solve complex non-linear problems of economic variables (Jain & Payal (2011)). LSTM has proven itself as a time-series analysis tool by solving vanishing gradient problems in Recurrent Neural Networks (RNN) and capturing long-term dependency. In this study, we built a LSTM model that can be used to forecast future value in various fields including stocks and real estates (Selvin et al. (2017); Lee and Jeon (2018); Bae and Yu (2018); Bi et al. (2019)).

Moreover, the Savitzky–Golay filter, which is used to remove signal noise in the signal processing field, is applied to the economic time-series variables to prove the superiority of the prediction performance after applying the denoising filter. Ultimately, the purpose of this research is to propose the optimal methodology in newbuilding ship price forecasting using the above process.

Objectives and Scope

In Chapter 2, a literature review introduces research and results for the validation of Efficient -Market Hypothesis (EMH) in the ship price market and is used to identify candidate groups for variables to be used in this study after examining the factors influencing ship price. Chapter 3 describes the methodologies of the VECM and LSTM; Savitzky–Golay filter, which is used to test hypotheses; and 34 years of experimental data. In Chapter 4, we first conduct a causality analysis to examine the effectiveness of economic variables identified as determinants affecting newbuilding ship prices in previous studies. For the VECM, a cointegration test is conducted in advance to identify the existence of long-term equilibrium between variables. As a result of the cointegration test, VECM, not VAR, is chosen to explain the forecasting model. Next, we evaluate the prediction performance after optimizing the LSTM model with hyper-parameters. In addition, we propose an improved model using Savitzky–Golay filter and describe the comparison between results and discussion. Lastly, Chapter 5 encapsulates the results of the above experiments, suggests the optimal time-series forecasting methodology in newbuilding ship price market, and summarizes the implications of this study.

2. Literature Review

2.1. EMH

Beenstock (1985) proposes a dynamic general equilibrium model to determine ship prices, assuming market efficiency and Rational Expectations (RE) in market. Because these assumptions imply that the predicted ship values by model are equivalent to the expected prices made by the market participants, his work triggered a number of subsequent studies, including Hale and Vanags (1992), Glen (1997), and Kavassanos and Alizadeh (2002), to validate EMH and RE. Strandenes (1984), meanwhile, assumes semi-RE in the price determination of second-hand ship prices. She explains that second-hand ship prices can be determined by the interest rates in current market and long-term equilibrium time charter rates, which represent the maritime transport market. Beenstock and Vergottis (1989a, b) formulate the ship price market using an asset pricing model. They argue that newbuilding and second-hand ships are strong substitutes, and prices between them are perfectly correlated. This approach is followed by Beenstock and Vergottis's (1992, 1993) subsequent works. This position, however, has been challenged by Haralambides et al. (2005) because second-hand ship prices are market driven, whereas newbuilding ship prices are supply and cost driven.

Hale and Vanags (1992) analyze second-hand ship prices of 30,000, 70,000, 120,000 dwt bulk carrier, and they cast doubt on the validity of the EMH and RE because of the evidence of the presence of cointegration in the dry cargo market.[1] Glen (1997) extends the expectations hypothesis validation in dry cargo and tanker market using Johansen's cointegration test. They comment that the existence of a cointegration relationship does not mean the failure of EMH, but that the market could be efficient because it is difficult to predict the price by stochastic characteristics in the long-term relationship. Kavussanos and Alizadeh (2002) investigate the validity of EMH and RE in ship price determination using VAR and Generalized Auto Regressive Conditional Heteroskedasticity (GARCH) models. Using data from 1976 to 1997, they conclude that the EMH in the second-hand and newbuilding markets are both rejected, and there exists a possibility of gaining excess returns on investment in ship price market.

[1] This opinion is refuted by Kavussanos and Alizadeh (2002), because the existence of cointegration is a sufficient condition, not a necessary condition for the efficient market.(Goulielmos (2019)).

2.2. Determinants of Ship Prices

As mentioned above, on the one hand, there have been many attempts to prove the existence of efficiency in ship price market using an econometric model, and, on the other hand, there has been much effort in analyzing the determinants of ship prices.

Veenstra (1999) conducts a cointegration test with newbuilding ship prices, second-hand ship prices, ship demolition prices, and freight rates. His work reveals that there are three cointegration vectors between the four variables, which indicate that there is a long-term equilibrium relationship between ship prices and freight rates empirically. Tsolakis et al. (2003), in their research of an econometric modelling in the second-hand ship price market, test cointegration with second-hand ship prices, newbuilding ship prices, freight rates, LIBOR (London Inter-Bank Offered Rates), ratio of order amount of ships, and available fleet. They find evidence that the variables that have a long-term equilibrium with second-hand ship prices are freight rates and LIBOR. Prior research substantiates the belief that freight rates are the main determinants of second-hand ship prices (Kim et al. (2013); Kim et al. (2014)).

Park (1998) highlights that prices of newbuilding vessels are primarily determined by freight rates, which represent the maritime transport market and second-hand ship prices. He concludes that fluctuations in the Japanese yen’s exchange rate and oil price have little effect on newbuilding ship prices.

2.3. Implications and Limitations of Previous Studies

Based on the limited previous studies related to ship prices, we can suggest the following implications. Firstly, the earlier studies of the efficiency in the ship market using econometric analysis show that there is a possibility implying market inefficiency in the ship price market. This means that profits can be realized through arbitrary trading strategies. Secondly, prices of newbuilding and second-hand ship prices should be considered as separate, since second-hand ships are available to be traded directly in the market, whereas new vessels cannot be traded until after the construction period.

Earlier works, however, hardly study the dominating factors that determine newbuilding ship prices. Moreover, existing determinants insufficient to explain newbuilding ship prices since the attempts listed to verify determinants of ship prices in Table 1 are mainly concentrated before 2005. Therefore, variables affecting newbuilding ship prices must to be reexamined.

Table 1. Previous studies of the econometric modelling in ship prices

Authors	Subject	Data
Strandenes (1984)	Second-hand ship price determinant	Time charter rates and second-hand ship prices of Panamax from 1968 to 1981
Beenstock and Vergottis (1989a)	Econometric modelling in tanker market	Freight rates, second-hand ship prices, scrap prices and newbuilding ship prices of tanker from 1950 to 1986
Beenstock and Vergottis (1989b)	Econometric modelling in dry cargo freight and shipping market	Freight rates, second-hand ship prices, scrap prices, ageing, deliveries and newbuilding ship prices of dry cargo from 1950 to 1985
Hale and Vanags (1992)	EMH	Second-hand ship prices from 1979 to 1988
Beenstock and Vergottis (1993)	Newbuilding ship prices modelling	Ship prices, risk premium from 1960 to 1985
Glen (1997)	EMH	Second-hand ship prices from 1979 to 1995
Kavussanos and Alizadeh (2002)	EMH	Second-hand and newbuilding ship prices, scrap prices and profits from 1976 to 1997

Park (1998)	Factor analysis of newbuilding ship prices	Freight rates, newbuilding and second-hand ship prices, exchange rate and oil price from 1976 to 1996
Tsolakis et al. (2003)	Second-hand ship prices modelling	Second-hand ship prices, newbuilding ship prices, freight rates, LIBOR, ratio of order amount of ships and available fleet from 1960 to 2001
Haralambides et al. (2005)	Newbuilding ship prices modelling	Newbuilding ship prices, freight rates, second-hand ship prices, shipbuilding costs, shipyard capacity, vessel orderbook from 1960 to 2001

2.4. Expected Contributions

Reexamination of Variable Effectiveness Toward Newbuilding Ship Prices

This article investigates causalities of explanatory variables in the newbuilding market, and NPI modelling with Granger's (1969) methodology. Clarkson's NPI is widely accepted and disseminated to comprehensively estimate the current newbuilding market. In this study, the NPI is chosen as an indicator to represent the newbuilding ship price market.

It can be seen from previous studies that freight rates, second-hand ship prices and interest rates must be verified as the major variables that affect newbuilding ship prices. It is essential to verify mutual correlations between related variables because ship prices are affected by many economic variables. By not only employing the Granger's causality test, but also taking Johansen's (1991) approach, this paper surveys cointegration associations to clarify the existence of long-term equilibrium among variables.

Discovering Superiority of LSTM Based on Neural Network

Financial data is difficult to forecast due to high volatility and noise. To solve this problem, there exist studies in freight market to predict financial markets through various deep learning models. The first researchers to forecast maritime business using non-linear neural network modelling were Li and Parsons (1997). They investigated monthly freight rates in the tanker market from January 1980 to October 1995 and demonstrated that neural network significantly outperforms Auto Regressive Moving Average (ARMA) time-series model in their work. They explained that because economic activities are rarely linear, neural network is more capable of dealing with non-linearity when analyzing time-series data in freight market. Lyridis et al. (2004) built an Artificial Neural Networks (ANN) model to forecast Very Large Crude/oil Carriers (VLCC) spot freight rates using data from October 1979 to December 2002. They commented that neural network is suitable for non-linear economic variables and can constitute decision-making tools with volatile time-series problems. However, studies that applied neural network models to time-series forecasting are insufficient for the newbuilding ship price market.

LSTM advanced from ANN has recently reported better performance in time-series prediction than traditional econometric models based on the linearity assumption. In a study that predicts real estate price index, Bae and Yu (2018) commented that when the market situation moves with a constant trend, both econometric time-series and neural network forecasting methodologies exhibit meaningful predictive power. However, in the case of rapid changes, the time-series model that assumes linearity is difficult to predict, whereas LSTM, which is robust to non-linear changes, provides a meaningful performance. Lee and Jeon (2018) conducted research that forecasted Seoul House Price Index by comparing the performances between econometric models and neural networks. In his study, he empirically confirmed the superiority of the neural network series models, including RNN and LSTM, over econometric time-series models such as VAR and VECM. He argues that this result confirms that housing price forecasting using neural network models is more useful for data with high volatility and non-linearity.

This work extends the scope of research to the newbuilding ship price market and empirically demonstrates the superiority of LSTM by comparing its performance with that of VECM, which represents the econometric forecasting modelling with the linearity assumption in NPI. For this reason, we argue that LSTM, which is relatively robust to non-linear modelling, outperforms econometric models with linear combinations between variables in periods of volatile newbuilding ship price market conditions.

Application of Denoising Filter

Another contribution of this research is the application of denoising filter to demonstrate an improved performance in the newbuilding ship price market. Lee and Oh (2019) improved performance of LSTM using the Savitzky–Golay filter to make predictions of Korea Composite Stock Price Index 200 (KOSPI200) futures index. Bi et al. (2019) experimented with deep neural network prediction models for user task time by analyzing 25,362,157 tasks collected from Google production compute clusters for 29 days. They used the Savitzky–Golay filter to eliminate noise in the original sequence, and the LSTM combined with the Savitzky–Golay filter outperformed the commonly used LSTM.

To propose an improved time-series forecasting methodology, we built prediction models by passing economic variables related to newbuilding ship prices through the Savitzky–Golay filter and empirically confirmed that models passed through denoising filter exhibit better performance.

3. Methodology and Data

3.1. VECM

Spurious regression problem will not pose an issue in multivariate time-series variables with differentiation if those are non-stationary (Granger and Newbold (1974)). When estimating a time-series model, however, variable differentiation can result in a loss of information about the long-term equilibrium relationships between economic variables, even though variables become stationary. A cointegration relationship exists between time-series variables if a linear combination of non-stationary variables satisfies stationarity. Therefore, VECM with error correction term added to VAR model can reasonably explain the relationship between variables when a cointegration relationship exists (Engle and Granger (1987)).

The vector time-series variable y_t, composed of integral variables $I(1)$, can be estimated by the following VAR model (Johansen and Juselius (1990)):

$$\Delta y_t = \Pi y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + \varepsilon_t \tag{1}$$

Where Δy_t is $(k \times 1)$ vector, Π, Γ_i signify $(k \times k)$ coefficient matrix, and $\varepsilon_t = (\varepsilon_{1t}, \ldots, \varepsilon_{kt})$ is a white noise vector.

If $r \leq k - 1$, meaning that the rank of Π is less than the number k of integrated time-series variables with $I(1)$, Eq. (1) can be converted to Eq. (2), since there exist matrices $(k \times r)$ α and β, which make $\Pi = \alpha\beta'$, $\beta' y_t$ satisfy stationarity.

$$\Delta y_t = \alpha\beta' y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + \varepsilon_t \tag{2}$$

Where β is a cointegration vector matrix $(k \times r)$, and α is an adjustment coefficients vector matrix $(k \times r)$ for adjusting the unbalanced error $\beta' y_t$ to satisfy the equilibrium relationship. Eq. (2) describes a restoring mechanism of the long-term equilibrium through a dynamic coordination process α, when $\beta' y_t$ is not a zero, meaning that a random shock causes a temporary departure from the long-term equilibrium.

3.2. Savitzky–Golay Filter

Savitzky and Golay (1964) used a least-squares polynomial approximation to propose a data smoothing method for clearing noise. Compared with the other low pass filters used for smoothing, which are generally defined in the frequency domain transformed from the time domain, the Savitzky–Golay filter has the advantage of time-series data filtering because it is designed in the time domain. Also, the filter shows high efficiency for preserving maximum, minimum, and peak point characteristics after removing noise.

A data sequence of $2m + 1$ samples centered at $n = 0$, the polynomial equation $p(n)$ with noise discarded can be given as follows (Bi et al. (2019)):

$$p(n) = \sum_{p=0}^{N} a_k n^k \qquad n \in [-m, m] \tag{3}$$

Where N is the degree of a polynomial, and a_k is the k th coefficient of a Savitzky–Golay fit function; n means data points, and M is the "half-width" of the approximation interval (Schafer (2011)).

We minimize the mean-squared approximation error for the data sequence samples centered on $n = 0$.

$$\varepsilon_N = \sum_{n=-M}^{M} (p(n) - x[n])^2 = \sum_{n=-M}^{M} \left(\sum_{p=0}^{N} a_k n^k - x[n] \right)^2 \tag{4}$$

3.3. LSTM

The traditional RNN is an algorithm used to learn data from a previous state for the current data learning. The output of the former state's hidden layer in RNN enters the cell as the input of the current hidden layer, which is computed with the current input layer to the output of the cell.

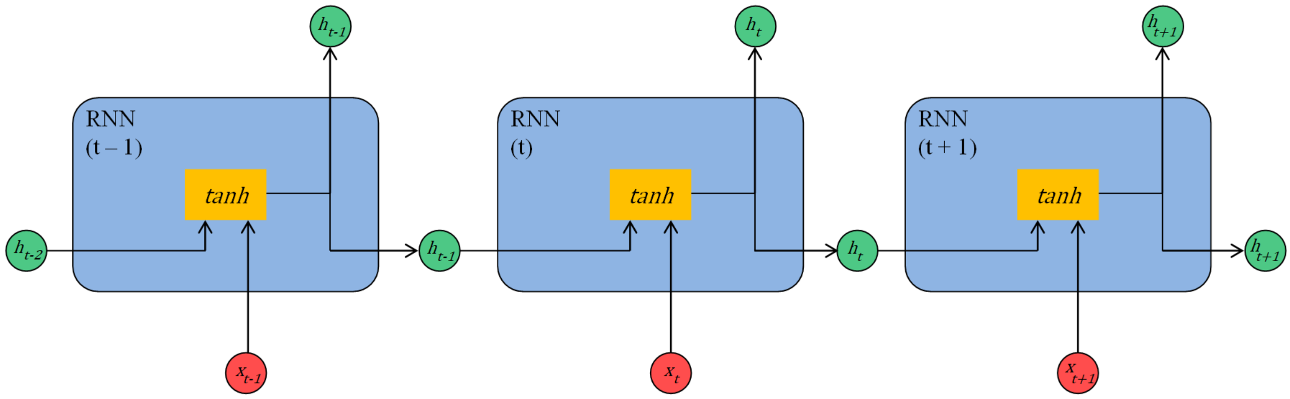


Figure 1. Structure of RNN

This network uses back-propagation to pass the error gradient back to each neuron for weight updates. As long as the distance between the time intervals increases, the vanishing gradient problem occurs whereby gradient values converge to zero (Pascanu et al. (2013)).

LSTM is a special structure of RNN and refers to a network that uses LSTM blocks as hidden layers. It has been proposed that it can overcome slow learning speed and vanishing gradient problems in RNN (Hochreiter and Schmidhuber (1997)). Figure 2 displays the structure of LSTM, which contain memory-moving cells that maintain the state over time as well as three non-linear gate structures that control data flow.

Denoted as f_t, the output of forget gate determines the amount of the information flow with a function of the current input values and the previously hidden layer output values:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \qquad f_t \in [0, 1] \tag{5}$$

Where $\sigma = \frac{1}{1+ e^{-x}}$ and a reflection ratio of the previous stage block value is determined by the product of the output of forget gate and the block state value of the previous stage.

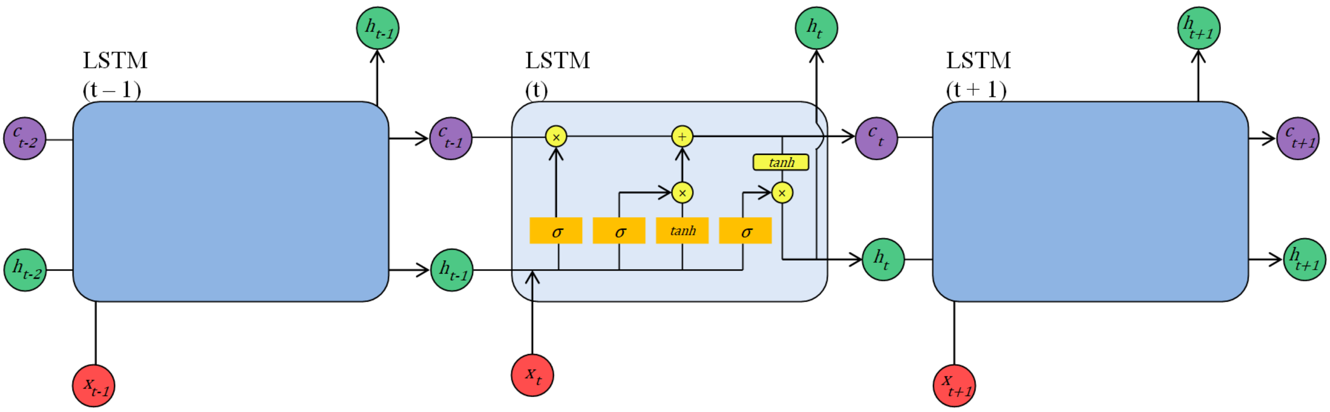


Figure 2. Structure of LSTM

As the following step illustrates, the input gate is a selection process whereby new information is added to the values passed through the forget gate. First, i_t, which determines a reflection ratio of new input values can be formulated as follows:

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \qquad i_t \in [0, 1] \tag{6}$$

The candidate value to be added to the new value denoted as $\widetilde{c}_t$ is determined by the *tanh* activation function layer of the output of the previously hidden layer and the current input. LSTM block state value c_t and $\widetilde{c}_t$ are set up as follows:

$$\widetilde{c}_t = tanh(W_c[h_{t-1}, x_t] + b_c) \qquad \widetilde{c}_t \in [-1, 1] \tag{7}$$

$$c_t = c_{t-1} \odot f_t + \widetilde{c}_t \odot i_t \tag{8}$$

The output gate eventually determines a reflection ratio o_t of a new output passing through sigmoid function σ and makes a new output candidate value h_t with passing c_t through the *tanh* activation function layer:

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \qquad o_t \in [0, 1] \tag{9}$$

$$h_t = tanh\,(c_t) \odot o_t \tag{10}$$

Where $[W_f, W_i, W_c, W_o]$ are the input weights, and $[b_f, b_i, b_c, b_o]$ denote bias of layers.

3.4. Data

Variable Selection and Source

As mentioned in Chapter 2.2, previous studies have identified the determinants of ship prices, which consider freight rates as an important factor in variations in ship prices. This is because, given a positive correlation between ship prices and freight rates, shipowners expect freight rates to create values for the future (Strandenes (1984); Beenstock and Vergottis (1989a, 1989b); Park (1998); Veenstra (1999); Haralambides et al. (2005)).

Another determinant explaining the change in ship prices is the interest rate. Tsolakis et al. (2003) link the interest rates to the ship prices as the interest expense to acquire expensive ship assets; they showed that interest rates can affect ship prices negatively.

Park (1998) disclosed that second-hand ship prices result in high levels of newbuilding orders and newbuilding ship prices can be predicted directly using a function of second-hand ship prices with a high correlation of 0.93. In addition, higher second-hand ship prices tend to increase the demand for new ships (Haralambides et al. (2005)).

In this research, NPI from Clarkson Research is selected as a comprehensive indicator of newbuilding ship prices. The index makes monthly comparisons with a ship construction cost of 100 as of January 1988, which means that newbuilding ship prices are higher as the index is greater than 100. And we use SPI to illustrate the comprehensive price of second-hand ships.

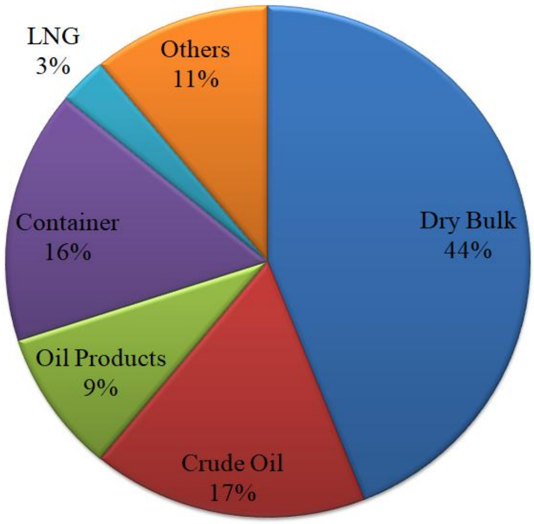


Figure 3. World Seaborne Trade in 2018

Source : Clarksons

The Baltic Shipping Exchange reports the Baltic Dry Index (BDI), which measures the average charter rate for bulk carriers transporting major cargoes, such as iron ore, coal, grain, etc. As shown in Figure 3, dry bulk holds the largest proportion of the world maritime trade. Therefore, BDI is used as a representative freight index of irregular shipping market and used to reflect freight rate index (Ahn and Lee (2018)).

Newbuilding ships in the market are generally traded in US dollars. LIBOR is the most widely used interest rates in ship financing, and it refers to the interest expense of funds borrowed when securing ships (Choi and Kim (2016)). Therefore, the US dollar three-month LIBOR rate that is used most representatively in the ship financial market was chosen.

Generally, second-hand ship prices serve as a leading indicator of ship demand (Karakitsos and Varnavides (2014)). An increase of seaborne trade will lead to an increase of the demand of shippers and shipowners to secure second-hand ships, and this promotes a rise of second-hand ship prices. As a result, the difference between second-hand ship prices and newbuilding ship prices decreases, and shipowners turn to newbuilding ship orders from second-hand ship market. For this reason, newbuilding ship prices are affected by second-hand ship prices, and we chose second-hand ship prices as a determining factor that suggests the direction of newbuilding ship prices.

The collected data contained information from every month beginning from January 1986 until July 2019. NPI and SPI were obtained from Clarkson Shipping Intelligent Network. BDI data from January 1986 to October 1999 were obtained from Korea Shipping Gazette and from Clarkson Shipping Intelligent Network from November 1999 to June 2019. The US dollar three-month LIBOR rate was collected from Economic Statistics System, Bank of Korea.

Table 2. Economic data types

Abbreviation	Full Name	Period	Source
NPI	Newbuilding Price Index	Jan 1986 – June 2019	Clarkson Shipping Intelligent Network
SPI	Second-hand Price Index		Clarkson Shipping Intelligent Network
BDI	Baltic Dry Index		Clarkson Shipping Intelligent Network, Korea Shipping Gazette
LIBOR3	US dollar three-month LIBOR rate		Economic Statistics System, Bank of Korea

Data Features

According to Stopford (2009), the shipping industry has frequent cycles, with each cycle repeating every eight years on average. However, no clear cyclical length is revealed, despite many attempts to forecast shipping cycles since Tinbergen (1931).

Three major depressions in the shipping industry since the Great Depression of 1929 are as follows (Goulielmos (2019)):

(1) The oil tankers crisis of 1973.

(2) The dry cargo depression (1981 to 1987).

(3) Lehman Brothers crisis (2008 to 2016).

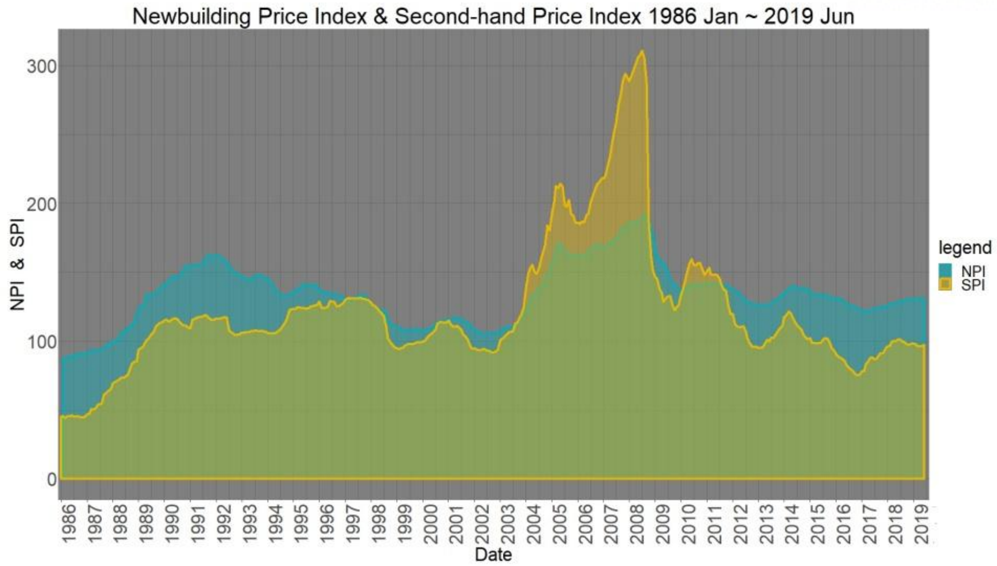


Figure 4. NPI and SPI (January 1986–June 2019)

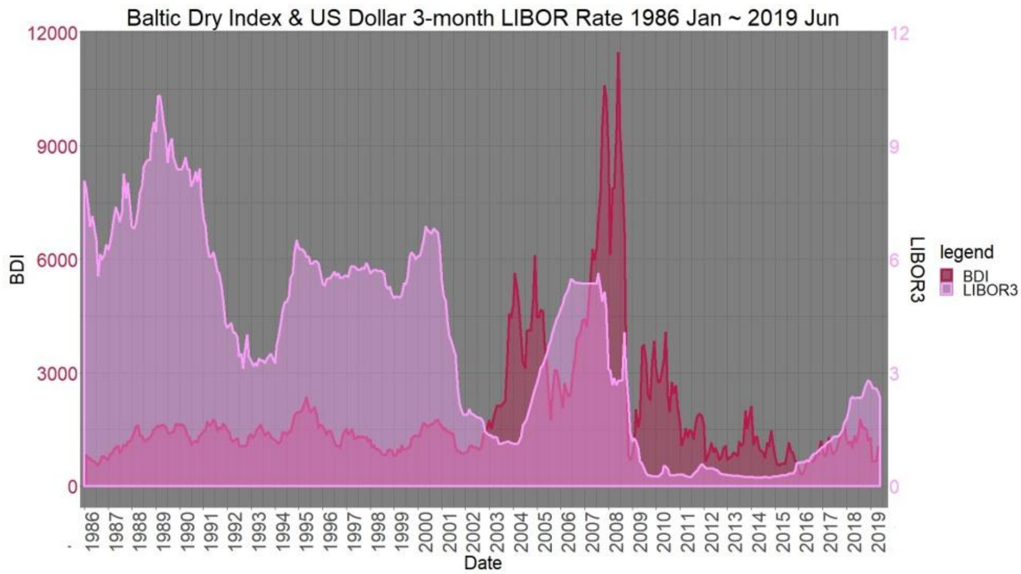


Figure 5. BDI and LIBOR3 (January 1986–June 2019)

All the economic variables covered in this study, including NPI, SPI, BDI, and LIBOR3, have been falling sharply since the Lehman Brothers crisis of 2008. This research includes the periods following the economic crisis of the Lehman Brothers of 2008. We would like to argue that the boom and recession have been offset within the cycle, although data have changed sharply since 2008 because the global shipping economy is an industry of iterative boom and recession cycles (Ahn and Lee (2018)).

Table 3. Descriptive statistics

Variable	NPI	SPI	BDI	LIBOR3
Average	133.67	122.13	1899	3.75
Median	133.38	110.63	1374	3.54
Max	191.51	310.90	11458	10.31
Min	87.09	44.10	314	0.22
Std. Dev	21.99	48.86	1685.48	2.72
Observations	402 Obs.			
Period	Jan 1986 ~ June 2019			

4. Empirical Results

4.1. Unit-Root Test

To ensure that spurious regression would not pose an issue for our analysis, we conducted the Augmented Dickey–Fuller (ADF) and Phillips–Perron (PP) unit-root tests to identify the stationarity of economic time-series variables.

Table 4. Results of unit-root tests

Variable	ADF		PP		Test Result
	Level	1st Difference	Level	1st Difference	
NPI	-2.822	-4.674^{**}	-1.897	-11.568^{**}	$I(1)$
SPI	-2.604	-5.609^{**}	-2.050	-13.003^{**}	$I(1)$
BDI	-2.935	-9.372^{**}	-3.224	-17.282^{**}	$I(1)$
LIBOR3	-3.478	-5.189^{**}	-1.942	-12.691^{**}	$I(1)$

** : Rejection of the hypothesis at the 0.01.

Results in Table 4 indicate that all level variables, including NPI, SPI, BDI, and LIBOR3, are non-stationary because the null hypothesis, "unit root is not present" is rejected from among our variables at 1% level. Therefore, four-time-series variables in our experiment yield $I(1)$ series because they show stationarity after first differentiation.

4.2. Granger Causality Test

When multivariate variables are used to estimate the time-series model, it is necessary to verify the causal relationship among the variables included. Granger (1969) proposed a test that provides information about the contribution of one variable to the prediction of another variable. The null hypothesis H_o of forecasting relationship between time-series variables X and Y is set to "X is not Granger Cause Y", and the rejection of H_o leads to the conclusion that "X Granger Causes Y". Table 5 shows the Granger causality test results between NPI and other economic variables.

Table 5. Granger causality test results between NPI, SPI, BDI, LIBOR3

Lag orders / H_o	1	2	3	4	5	6	7	8	9	10
SPI↛NPI	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$
BDI↛NPI	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$
LIBOR3↛NPI	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$
NPI↛SPI	0.030^{*}	0.049^{*}	0.158	0.259	0.350	0.457	0.581	0.730	0.690	0.526
BDI↛SPI	$< 0.001^{***}$	0.009^{**}	0.031^{*}	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$
LIBOR3↛SPI	0.227	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$
NPI↛BDI	0.895	0.092	0.065	0.067	0.016^{*}	0.036^{*}	0.049^{*}	0.087	0.179	0.166
SPI↛BDI	0.092	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$
LIBOR3↛BDI	0.850	0.952	0.131	0.178	0.069	0.046^{*}	0.050	0.060	0.112	$< 0.001^{***}$
NPI↛LIBOR3	0.040^{*}	0.026^{*}	0.064	0.098	0.239	0.175	0.141	0.027^{*}	0.047^{*}	0.155
SPI↛LIBOR3	0.278	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$	$< 0.001^{***}$	0.001^{**}
BDI↛LIBOR3	0.353	0.040^{*}	0.031^{*}	0.001^{**}	0.003^{**}	0.004^{**}	0.006^{**}	0.004^{**}	0.006^{**}	0.015^{*}

***, **, * : H_o is rejected under 0.1%, 1%, 5% significance level, respectively.

As shown in Table 5, the results demonstrate that SPI, BDI, and LIBOR3 Granger cause NPI to have a significance level of less than 0.1%. This strongly supports previous studies that second-hand ship prices alongside BDI and LIBOR are factors that determine newbuilding ship prices.

In the causality test among the economic variables, including SPI, BDI, and LIBOR rates, there are significant bidirectional causal relationships at 0.1% or 5% significance level under lags in 10 months.

4.3. Performance Measurement and Validation

Performance Measurement

In this chapter, we use loss functions to evaluate the performance of forecasting methodologies. Root Mean Square Error (RMSE) is a measurement based on the square of the distance between real and predicted values. Hence,

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y_t})^2} \tag{11}$$

Mean Absolute Percentage Error (MAPE) is a scale-independent measurement of the difference between real and predicted values, expressed as a percentage. Hence,

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{y_t - \hat{y_t}}{y_t}\right| \times 100 \tag{12}$$

Where n is the number of the instances of monthly data, and y_t and $\hat{y_t}$ denote the observed and predicted values at time t.

Validation : Sliding Window Test

We conducted a sliding window test that moves with the set periods of test dividing time intervals. Figure 6 shows a sliding window test used to measure the robustness of predicted models in time-series forecasting.

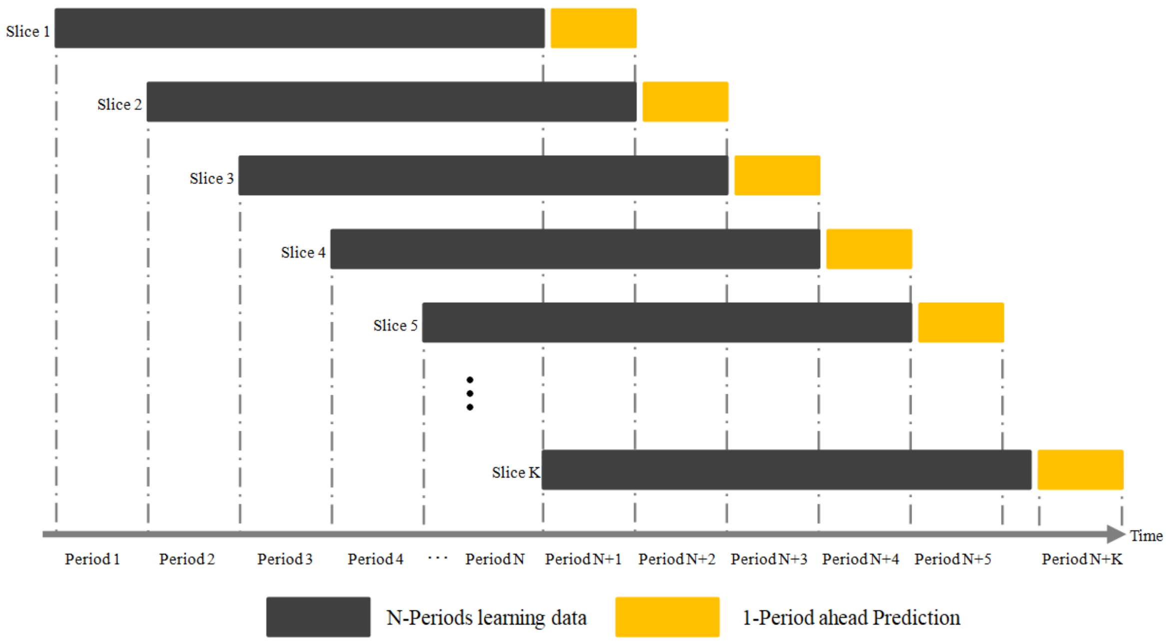


Figure 6. Sliding window test

We choose this validation framework in line with Jo et al. (2018) and Kim and Oh (2019). The sliding window test in time-series forecasting is useful because it reflects a progressive reduction in the impact of historical data by not using all datasets for model training. Also, a sliding window test consists of sequential data divided into windows with multiple overlapping periods that repeat training and testing.

4.4. VECM Modelling

Cointegration test

We first conducted a log-transformation on the variables with different scale levels, including NPI, SPI, and BDI. These log-transformed variables are used to calculate the VECM model. Using the cointegration test, a multivariate VAR model can be extended to the VECM model with the equilibrium error if cointegration relationships exist between variables that follow the non-stationary process.

The lag order should be determined first in the Johansen cointegration test since the error correction model is estimated in the VECM model in Eq. (1) and Eq. (2). We used information criteria such as Akaike Information Criterion (AIC), Schwarz Information Criterion (SIC), and Hannah-Quinn Information Criterion (HQC), to determine the optimal lag length in the VECM model.

Table 6. Optimal lag length of the VECM model

Optimal lag order	NPI	SPI	BDI	LIBOR3	NPI + SPI + BDI + LIBOR3
AIC	7	3	5	10	4
SIC	4	3	2	2	2
HQC	4	3	2	10	2

* Test period : 1 ~ 15 months

The test results are shown in Table 6. We used a lag length of 2 since we followed the principle of parsimony to choose the most straightforward lag order of model among AIC of 4 order, SIC of 2, and HQC of 2.

To examine the existence of long-term equilibrium between economic variables, we used Johansen's (1991) approach. The Johansen cointegration test follows the process of estimating and testing the number of cointegration relationships and the coefficients of a model with Maximum Likelihood Estimation (MLE) instead of the Engle and Granger's (1987) approach, which can yield different estimations according to the order of test variables.

Table 7. Johansen cointegration rank test

H_o : Rank=r	H_a : Rank>r	Eigenvalue	Trace	0.05 Critical value
0	0	0.18678	102.51	62.99
1	1	0.02949	20.35	42.44
2	2	0.01753	11.06	25.32
3	3	0.00992	3.99	12.25

The results reported in Table 7 show evidence of existence of cointegration between the variables. First, the null hypothesis "H_o : Rank = 0" is rejected. Second, "H_o : Rank = 1" cannot be rejected since the critical value is 42.44 below the significance level of 5%, whereas the trace statistic value is 20.35. In other words, a long-term equilibrium relationship exists between NPI, SPI, BDI, and LIBOR3. Therefore, it is reasonable to use the VECM model extended from the multivariate VAR model.

Results

VECM is a model set on the premise of cointegration. Given the existence of cointegration, we used the VECM model estimation using MLE.

Table 8. Estimates of the VECM model

	ΔNPI_t	
	Coefficient	Standard Error
Intercept	0.0004	0.0005
ΔNPI_{t-1}	0.2449	0.0496
ΔNPI_{t-2}	0.1902	0.0480
ΔSPI_{t-1}	0.0891	0.0184
ΔSPI_{t-2}	0.0383	0.0192
ΔBDI_{t-1}	-0.0031	0.0030
ΔBDI_{t-2}	0.0017	0.0030
$\Delta LIBOR3_{t-1}$	-0.0007	0.0021
$\Delta LIBOR3_{t-2}$	0.0055	0.0021
ECT (Error Correction Term)	-0.0016	0.0022

Table 9 shows the results of NPI monthly forecasts and draws a comparison between the predicted and actual values using four economic variables, namely NPI, SPI, BDI, and LIBOR3, from January 1986 to June 2019. The sliding window test is designed to validate models for their robustness, trained from 24 years of data shifted by one year and verified as an average predicted the value of test data from 2010 to 2019 each year.[2] Finally, the average performance of the VECM model that we experimented shows that RMSE is 3.47231 and MAPE is 0.02355.

[2] Test data in 2019 is from January to June.

Table 9. VECM forecasting using the sliding window test

Period		RMSE	MAPE
Training	Validation		
1986 Jan – 2009 Dec	2010 Jan – 2010 Dec	1.13142	0.00721
1987 Jan – 2010 Dec	2011 Jan – 2011 Dec	3.15776	0.02130
1988 Jan – 2011 Dec	2012 Jan – 2012 Dec	2.50323	0.01577
1989 Jan – 2012 Dec	2013 Jan – 2013 Dec	6.30639	0.03915
1990 Jan – 2013 Dec	2014 Jan – 2014 Dec	2.60661	0.01744
1991 Jan – 2014 Dec	2015 Jan – 2015 Dec	5.81748	0.04226
1992 Jan – 2015 Dec	2016 Jan – 2016 Dec	4.57660	0.03511
1993 Jan – 2016 Dec	2017 Jan – 2017 Dec	2.90184	0.01944
1994 Jan – 2017 Dec	2018 Jan – 2018 Dec	2.30291	0.01500
1995 Jan – 2018 Dec	2019 Jan – 2019 Jun	3.41914	0.02284
Avg.		3.47231	0.02355

4.5. LSTM Modelling

This chapter concerns the NPI prediction by designing the LSTM model with optimal hyper-parameters. As shown in the previous chapter, economic variables, including NPI, SPI, BDI, and LIBOR3 from January 1986 to June 2019, were used to estimate NPI. The sliding window test was used to compare and analyze the annual forecasting performance from 2010 to 2019.

Network Design

In the multivariate LSTM analysis, the economic variables are generally scaled through normalization (Li and Parsons (1997)). In this study, we converted each variable to Z-score normalization.

$$z_k = \frac{x_k - \mu}{\sigma} \tag{13}$$

Where μ is the arithmetic mean and σ is the standard deviation of the given data.

Depending on the situation, a proper hyper-parameter tuning in LSTM is the most important factor in building a trustworthy model, and inapposite parameter settings that ignore business issues can lead to under- or overfitting problems. An optimal neural network model such as ANN, RNN, and LSTM is generated by trial and error strategy (Lenk et al. (1997)). In other words, in no way can the *best* hyper-parameters in the LSTM model design be found. Nonetheless, it is necessary to create an *optimal* design through many trials and errors.

Furthermore, the goal of machine learning is to generalize the notion that the performance of a fitted model yields good results not only in training data, but also in other related data, meaning the optimal performance with test data input (Geron (2017)).

Drawing on the monthly data collected, we tried several steps to search for hyper-parameters of LSTM in moving years from 2010 to 2019, which, using the sliding window test, were set as test data, each trained from past 24 years data and derived from the optimal parameters, as elaborated below.

Step 1. Number of layers and cells

First, we conducted experiments with the optimal number of LSTM layers. The LSTM layer is a hyper-parameter that uses the number of hidden layers in the LSTM network. Second, the parameter of architecture constitutes the number of cells in each layer. In this test, the number of layers and cells of each layer is set to 4, 3, 2, 1 and 150, 120, 90, 60, 30, respectively. The experiment results show that the optimal number of layers is 3, and the number of cells of layers 1, 2, and 3 is 60, 60, and 60, respectively (see Figure 7).

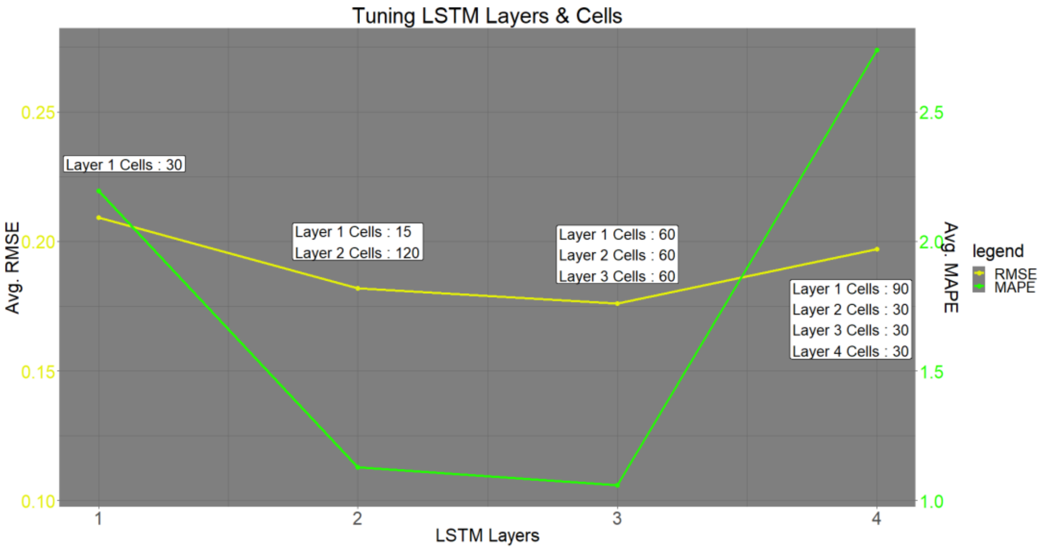


Figure 7. Tuning LSTM layers and cells

Step 2. Sequence length

LSTM has a unique loop structure that uses past information to affect a specific data point, which is useful for forecasting time-series data by exploring an adequate sequence length of data. At an arbitrary point in time t, the output y_t of the LSTM cell is a function of all inputs x_t, including y_{t-1} of the previous cell. Therefore, it is affected by the recurrent structure connected by the sequence length of LSTM.

Figure 8 shows the search results for an optimal sequence length of LSTM, which we found to be sequence length 4, minimizing RMSE and MAPE.

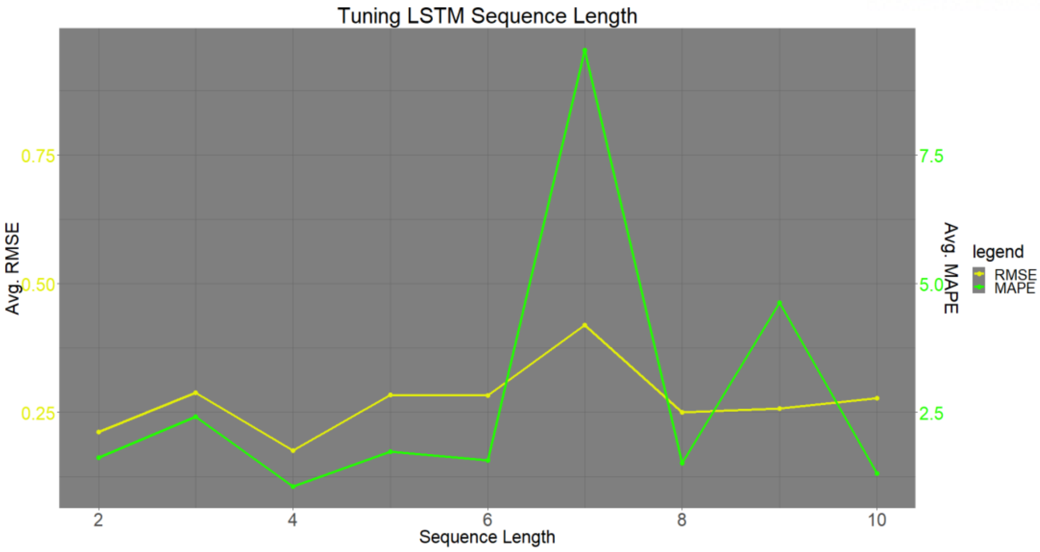


Figure 8. Tuning LSTM sequence length

Step 3. Activation functions of layers

As mentioned in Chapter 3.3, the candidate value to be added to new values is calculated as a function of the output of the previously hidden layer, the current input, and the activation function. We tuned activation functions, such as Rectified Linear Unit (ReLU), Tanh, Softmax, and Sigmoid, on each layer and found ReLU, Tanh, and Softmax on layers 1, 2, 3, respectively.

Step 4. Optimizers of layers

Optimizers adjust and update the weight parameters to minimize the loss function. They also contribute to the exploration of the optimal LSTM model. Table 10 displays the results of hyper-parameter tuning among the Adaptive with moment (Adam), Root Mean-Squared Prop (RMSProp), and Stochastic Gradient Descent (SGD) optimizers.

Table 10. Results of tuning optimizers

Optimizer	RMSE	MAPE
Adam	0.14816	1.73885
RMSProp	0.13731	1.57747
SGD	0.32032	2.69442

Step 5. Epoch

Epoch is a parameter that determines the number of times repeated learning is performed. On the one hand, the iterative learning tends to increase a model's learning performance; on the other hand, it has the possibility of falling into overfitting.

In this work, we tested epochs ranging from 1 to 1000 and found an elbow point of loss functions. Figure 9 shows RMSE and MAPE change as the epoch varies. This result indicates that epoch 150 is a reasonable choice because the loss function decreases abruptly before epoch 150, and after that point, the change of loss function is no longer significant, considering the trade-off relationship of performance.

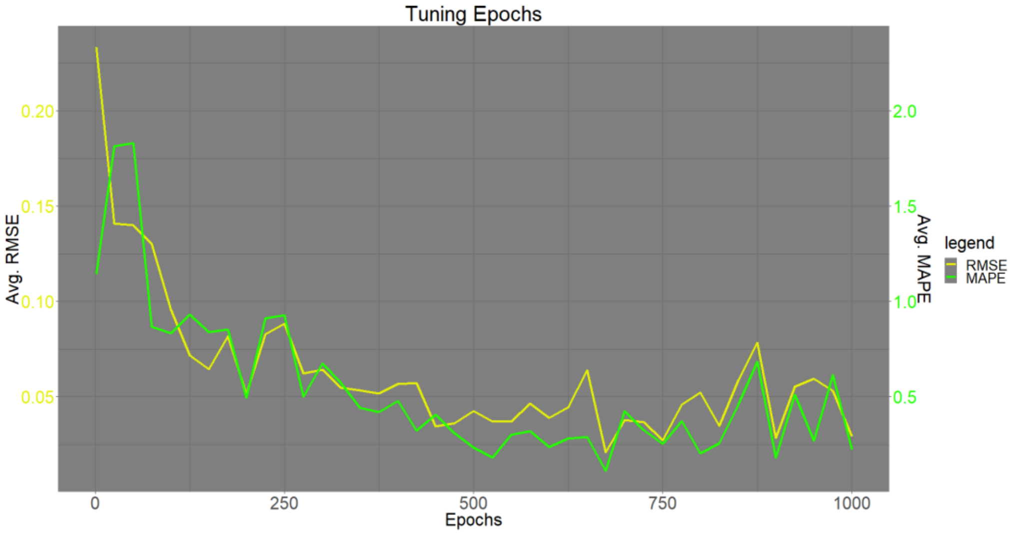


Figure 9. Loss function variation with epochs

Table 11 summarizes the hyper-parameters tuned.

Table 11. Hyper-parameter selection of LSTM architecture(v : variable)

Hyper-parameter	Value	Selected	
Number of Layers	[1, 2, 3, 4]	3	
Number of Cells	[30, 60, 90, 120, 150]	Layer 1	60
		Layer 2	60
		Layer 3	60
Sequence Length	[1 - 10] by 1	4	
Activation Function	[ReLU, Tanh, Softmax, Sigmoid]	Layer 1	ReLU
		Layer 2	Tanh
		Layer 3	Softmax
Optimizer	[Adam, RMSProp, SGD]	RMSProp	
Epoch	[1 - 1000] by 25	150	

Results

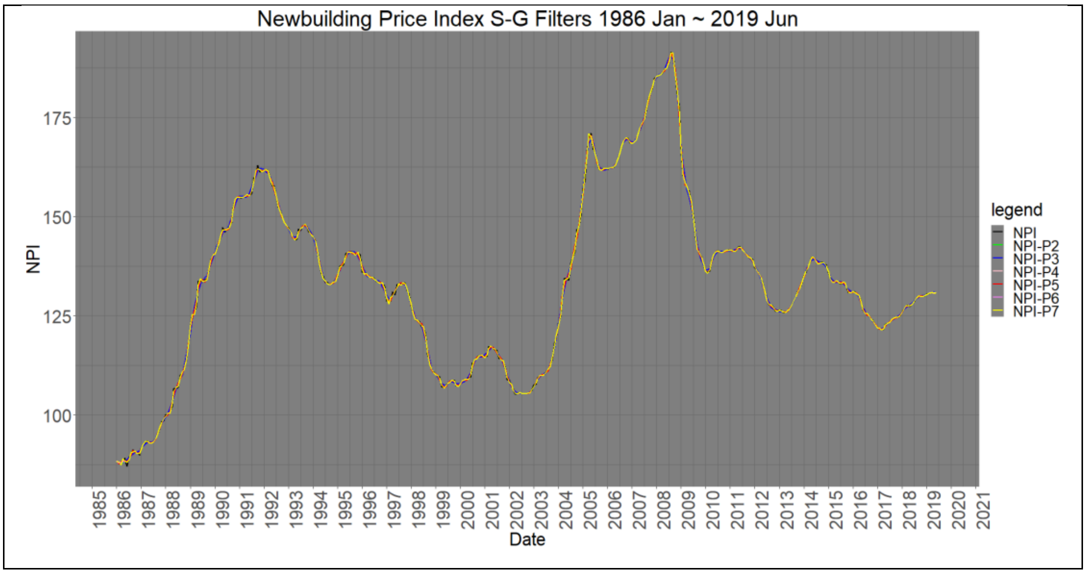
Similar to Chapter 4.4, we fitted the LSTM model that estimates NPI with data from January 1986 to June 2019. The sliding window method was used to validate the model between the average predicted and actual values of each year from 2010 to 2019 with a fitted model over the past 24 years. The average RMSE and MAPE of models are 1.41537, and 0.00981, respectively (see Table 12).

Table 12. LSTM forecasting using sliding window test

Period		RMSE	MAPE
Training	Validation		
1986 Jan – 2009 Dec	2010 Jan – 2010 Dec	1.47252	0.01049
1987 Jan – 2010 Dec	2011 Jan – 2011 Dec	0.83923	0.00594
1988 Jan – 2011 Dec	2012 Jan – 2012 Dec	2.70317	0.01930
1989 Jan – 2012 Dec	2013 Jan – 2013 Dec	0.77796	0.00530
1990 Jan – 2013 Dec	2014 Jan – 2014 Dec	0.76447	0.00415
1991 Jan – 2014 Dec	2015 Jan – 2015 Dec	1.88179	0.01383
1992 Jan – 2015 Dec	2016 Jan – 2016 Dec	2.47357	0.01435
1993 Jan – 2016 Dec	2017 Jan – 2017 Dec	0.36364	0.00253
1994 Jan – 2017 Dec	2018 Jan – 2018 Dec	2.75772	0.02143
1995 Jan – 2018 Dec	2019 Jan – 2019 Jun	0.11962	0.00079
Avg.		1.41537	0.00981

4.6. Denoise Filter

In this chapter, we suggest the denoising filter for improving the forecasting performance of VECM and LSTM by Savitzky–Golay filter, which minimizes data damage and removes noise by using the method of least squares. Figure 10 shows the denoising smoothing effect with various polynomial orders from January 1986 to June 2019.



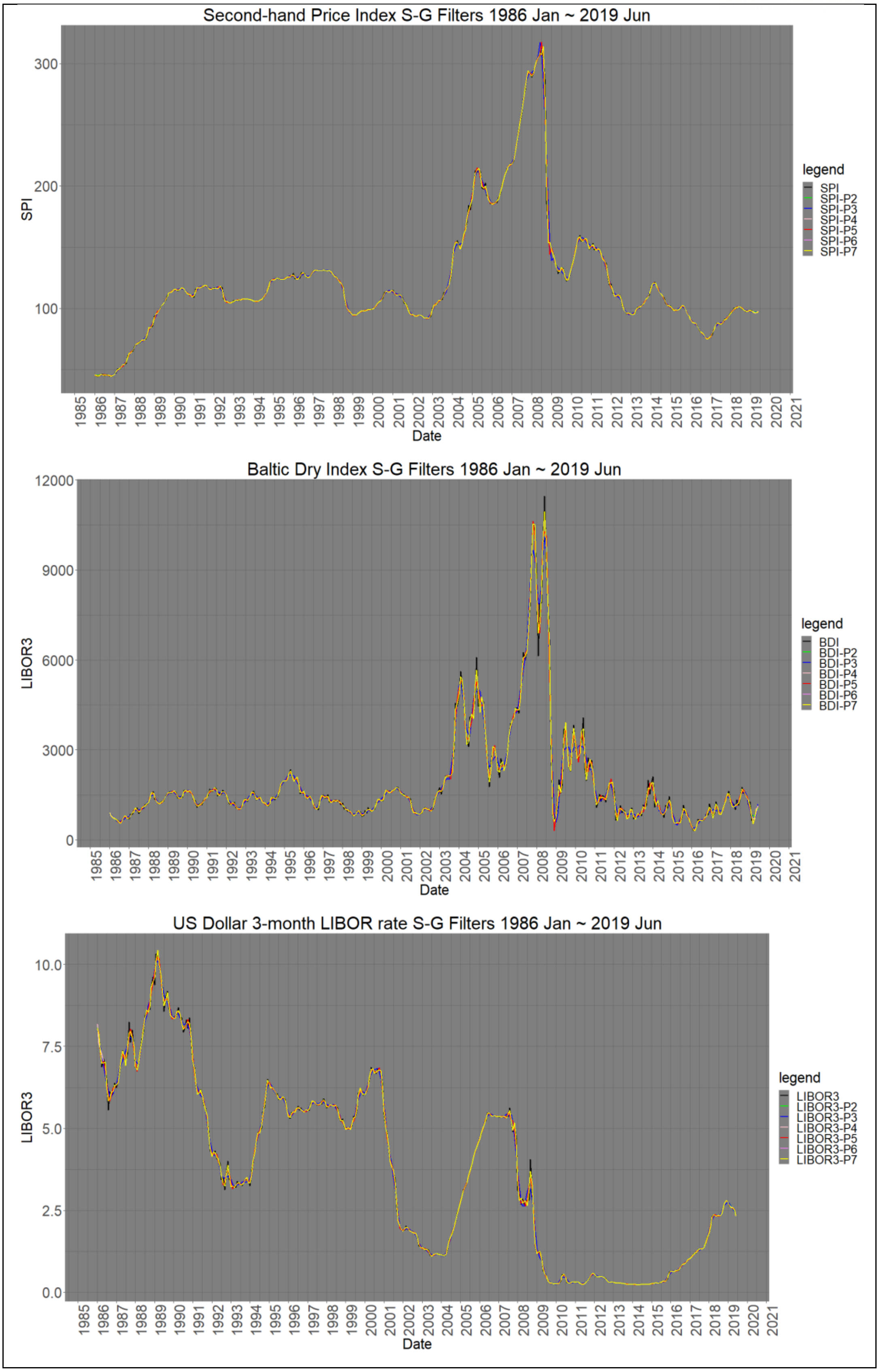


Figure 10. NPI, SPI, BDI, and LIBOR3 of Savitzky–Golay filters for window size 9 and various polynomial orders

This polynomial filter was used to remove noise to predict better VECM and LSTM models built in Chapters 4.4 and 4.5, respectively; the comparison results will focus on the preceding and subsequent stages of denoise filtering. By applying to economic variables different polynomial orders ranging from 2 to 6 and window sizes below 15, including NPI, SPI, BDI, and LIBOR3, we will further explore the parameters that optimize the predictive performance of NPI.

Savitzky–Golay VECM (SG–VECM)

Table 13 shows the results of loss functions, including RMSE and MAPE, by applying different polynomial orders and window sizes of the Savitzky–Golay filter for noise smoothing.[3] It shows an optimal denoising result when predicting SG–VECM ($p = 3$ and $n = 15$).

Table 13. Parameter settings for SG–VECM

Polynomial Order(p)	Window Size(n)	Loss	
		RMSE	MAPE
2	3	3.43822	0.02331
	5	2.78470	0.01863
	7	2.50388	0.01674
	9	2.88295	0.01871
	11	2.23195	0.01455
	13	2.20800	0.01414
	15	2.17793	0.01448
3	5	2.78222	0.01861
	7	2.49996	0.01671
	9	2.94053	0.01907
	11	2.25693	0.01471
	13	2.22640	0.01425
	15	**2.16390**	**0.01441**
4	5	3.43822	0.02331
	7	2.98320	0.01978
	9	2.80132	0.01860
	11	2.88597	0.01885

[3] The VECM model is identical to the model in Table 8.

4	13	2.70369	0.01731
	15	2.54477	0.01648
5	7	2.98980	0.01985
	9	2.79835	0.01852
	11	2.93149	0.01915
	13	2.74530	0.01757
	15	2.53683	0.01644
6	7	3.43822	0.02331
	9	3.14698	0.02086
	11	2.86462	0.01898
	13	2.79974	0.01867
	15	2.85489	0.1839

We observed a better performance with the average RMSE being 2.16390 and MAPE 0.01441, forecast by SG–VECM, and compared with the VECM model over the same period when RMSE was 3.47231 and MAPE was 0.02355 (see Table 14).

Table 14. SG-VECM forecasting using sliding window test with $p = 3$, and $n = 15$

Period		RMSE	MAPE
Training	Validation		
1986 Jan – 2009 Dec	2010 Jan – 2010 Dec	3.68511	0.02471
1987 Jan – 2010 Dec	2011 Jan – 2011 Dec	2.00346	0.01230
1988 Jan – 2011 Dec	2012 Jan – 2012 Dec	3.22873	0.02342
1989 Jan – 2012 Dec	2013 Jan – 2013 Dec	3.60415	0.01846
1990 Jan – 2013 Dec	2014 Jan – 2014 Dec	1.52655	0.00989
1991 Jan – 2014 Dec	2015 Jan – 2015 Dec	1.79197	0.01332
1992 Jan – 2015 Dec	2016 Jan – 2016 Dec	2.82730	0.02205
1993 Jan – 2016 Dec	2017 Jan – 2017 Dec	0.77902	0.00445
1994 Jan – 2017 Dec	2018 Jan – 2018 Dec	1.29817	0.01006
1995 Jan – 2018 Dec	2019 Jan – 2019 Jun	0.89457	0.00545
Avg.		2.16390	0.01441

Savitzky–Golay LSTM (SG–LSTM)

Next we tested the Savitzky–Golay filter on the LSTM model, which was built in Chapter 4.5 to improve forecasting performance. Table 15 shows the results of loss functions with various tuned parameters of SG–LSTM. It also shows the best efficiency with $p = 6$ and $n = 13$.

Table 15. Parameter settings for SG–LSTM

Polynomial Order(p)	Window Size(n)	Loss	
		RMSE	MAPE
2	3	1.46011	0.01017
	5	1.95564	0.01337
	7	1.63024	0.01076
	9	1.30934	0.00907
	11	1.60305	0.01143
	13	1.58291	0.01155
	15	1.48009	0.01067
3	5	1.93525	0.01318
	7	1.60219	0.01045
	9	1.30511	0.00907
	11	1.68537	0.01196
	13	1.62462	0.01178
	15	1.67739	0.01217
4	5	1.50184	0.01041
	7	1.23360	0.00811
	9	1.73847	0.01231
	11	1.31578	0.00937
	13	1.72075	0.01218
	15	1.29210	0.00953
5	7	1.26568	0.00832
	9	1.72191	0.01219
	11	1.45182	0.01036
	13	1.53383	0.01082
	15	1.39173	0.01019
6	7	1.53025	0.01070

6	9	1.67981	0.01228
	11	1.71511	0.01239
	13	**0.57556**	**0.00389**
	15	1.29537	0.938

Based on Table 16, which shows the average RMSE to be 0.57556 and MAPE to be 0.00389, we found the evidence that, compared with LSTM, SG–LSTM shows superior performance when forecasting NPI in the ship price market.

Table 16. SG–LSTM forecasting using sliding window test with $p = 6$, and $n = 13$

Period		RMSE	MAPE
Training	Validation		
1986 Jan – 2009 Dec	2010 Jan – 2010 Dec	0.19767	0.00134
1987 Jan – 2010 Dec	2011 Jan – 2011 Dec	1.06254	0.00753
1988 Jan – 2011 Dec	2012 Jan – 2012 Dec	0.88192	0.00579
1989 Jan – 2012 Dec	2013 Jan – 2013 Dec	0.95269	0.00703
1990 Jan – 2013 Dec	2014 Jan – 2014 Dec	0.42267	0.00289
1991 Jan – 2014 Dec	2015 Jan – 2015 Dec	0.07225	0.00048
1992 Jan – 2015 Dec	2016 Jan – 2016 Dec	1.25947	0.00736
1993 Jan – 2016 Dec	2017 Jan – 2017 Dec	0.10517	0.00068
1994 Jan – 2017 Dec	2018 Jan – 2018 Dec	0.40995	0.00283
1995 Jan – 2018 Dec	2019 Jan – 2019 Jun	0.39127	0.00297
Avg.		0.57556	0.00389

4.7. Discussion

Table 17 compares RMSE and MAPE forecasting models fitted from January 1986 to June 2019.

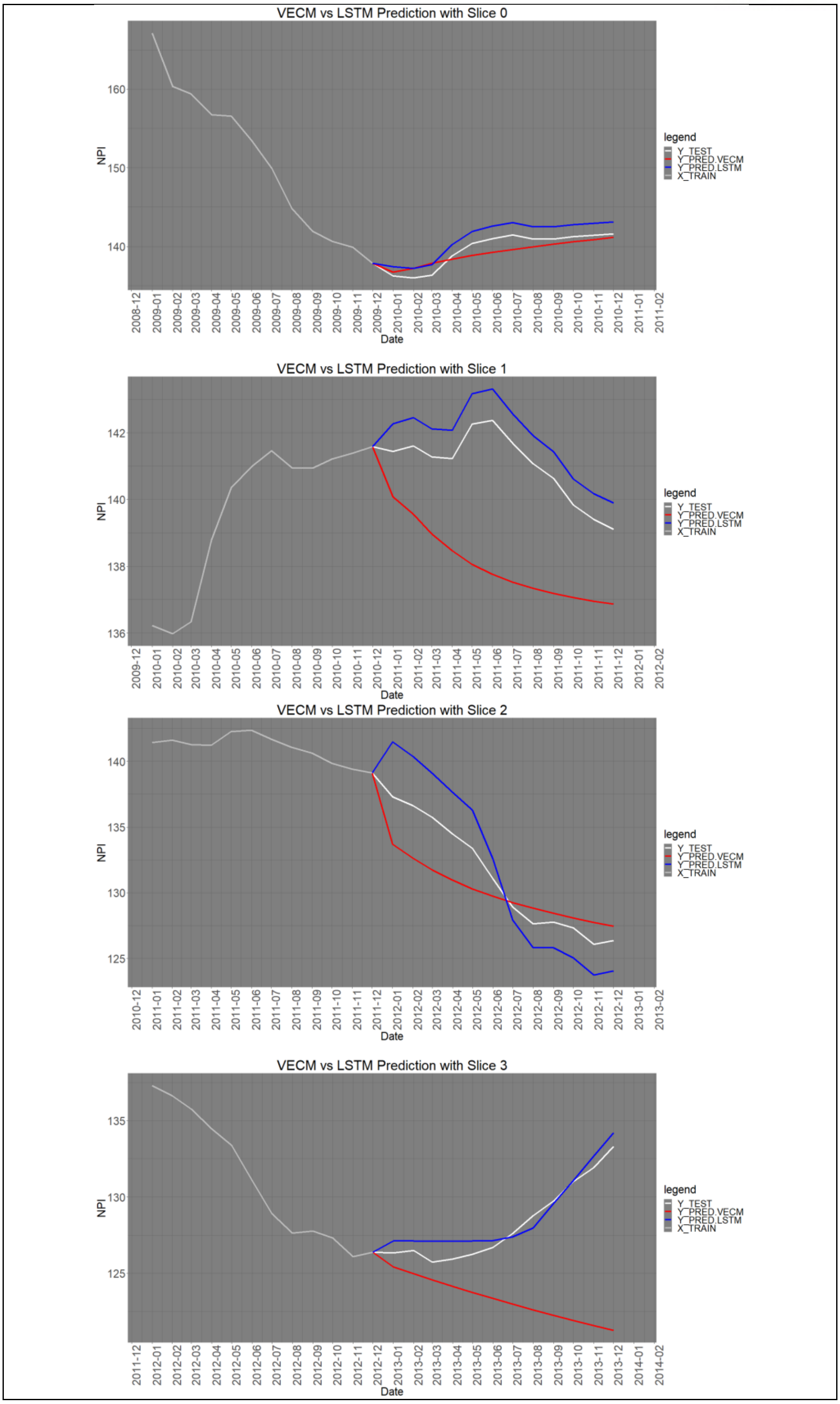
Table 17. Performance comparison from 2010 to 2019

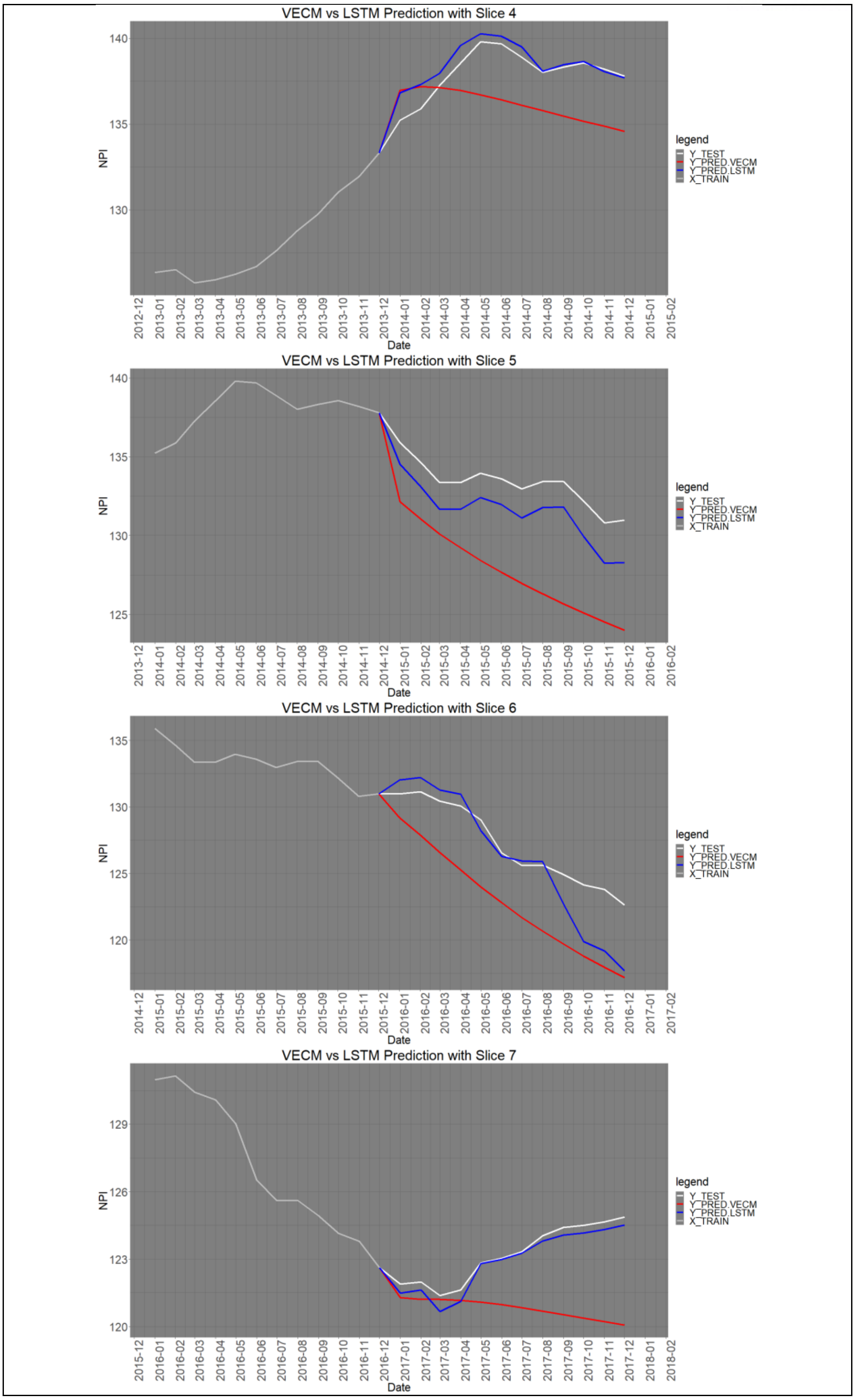
Model		2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	Avg.
VECM	RMSE	1.1314	3.1578	2.5032	6.3064	2.6066	5.8175	4.5766	2.9018	2.3029	3.4191	3.4723
	MAPE	0.0072	0.0213	0.0158	0.0392	0.0174	0.0423	0.0351	0.0194	0.0150	0.0228	0.0236
LSTM	RMSE	1.4725	0.8392	2.7032	0.7780	0.7645	1.8818	2.4736	0.3636	2.7577	0.1196	1.4154
	MAPE	0.0105	0.0059	0.0193	0.0053	0.0042	0.0138	0.0144	0.0025	0.0214	0.0008	0.0098
SG-VECM	RMSE	3.6851	2.0035	3.2287	3.6042	1.5266	1.7920	2.8273	0.7790	1.2982	0.8946	2.1639
	MAPE	0.0247	0.0123	0.0234	0.0185	0.0099	0.0133	0.0221	0.0045	0.0101	0.0055	0.0144
SG-LSTM	RMSE	0.1977	1.0625	0.8819	0.9527	0.4227	0.0723	1.2595	0.1052	0.4100	0.3913	0.5756
	MAPE	0.0013	0.0075	0.0058	0.0070	0.0029	0.0005	0.0074	0.0007	0.0028	0.0030	0.0039

The LSTM shows better performance when comparing the prediction results of VECM and LSTM. Figure 11 shows detailed results for the period set by test data from 2010 to 2019. They suggest that both VECM and LSTM show good performance when the market has a stable. However, in the case of volatile markets (e.g., Slice 3 in 2013 and Slice 7 in 2017), VECM hinders prediction while LSTM yields similar predictions to original test data.

This is because the VECM model assumes a linear fashion, which, in turn, reinforces Bae & Yu’s (2018) and Lee & Jeon’s (2019) arguments that machine learning models in time-series forecasting can also be used to predict market trends for volatile markets since they constitute robust to non-linear time-series modelling.

As for denoising filter, VECM and LSTM performances based on RMSE improved by 38% and 59% after applying the Savitzky–Golay filter. Finally, SG–LSTM shows the best result, with RMSE being 0.5756, and MAPE being 0.0039.


VECM vs LSTM Prediction with Slice 0
NPI
Date
legend
Y_TEST
Y_PRED.VECM
Y_PRED.LSTM
X_TRAIN
VECM vs LSTM Prediction with Slice 1
NPI
Date
legend
Y_TEST
Y_PRED.VECM
Y_PRED.LSTM
X_TRAIN
VECM vs LSTM Prediction with Slice 2
NPI
Date
legend
Y_TEST
Y_PRED.VECM
Y_PRED.LSTM
X_TRAIN
VECM vs LSTM Prediction with Slice 3
NPI
Date
legend
Y_TEST
Y_PRED.VECM
Y_PRED.LSTM
X_TRAIN


VECM vs LSTM Prediction with Slice 4
NPI
Date
legend
Y_TEST
Y_PRED.VECM
Y_PRED.LSTM
X_TRAIN
VECM vs LSTM Prediction with Slice 5
NPI
Date
legend
Y_TEST
Y_PRED.VECM
Y_PRED.LSTM
X_TRAIN
VECM vs LSTM Prediction with Slice 6
NPI
Date
legend
Y_TEST
Y_PRED.VECM
Y_PRED.LSTM
X_TRAIN
VECM vs LSTM Prediction with Slice 7
NPI
Date
legend
Y_TEST
Y_PRED.VECM
Y_PRED.LSTM
X_TRAIN

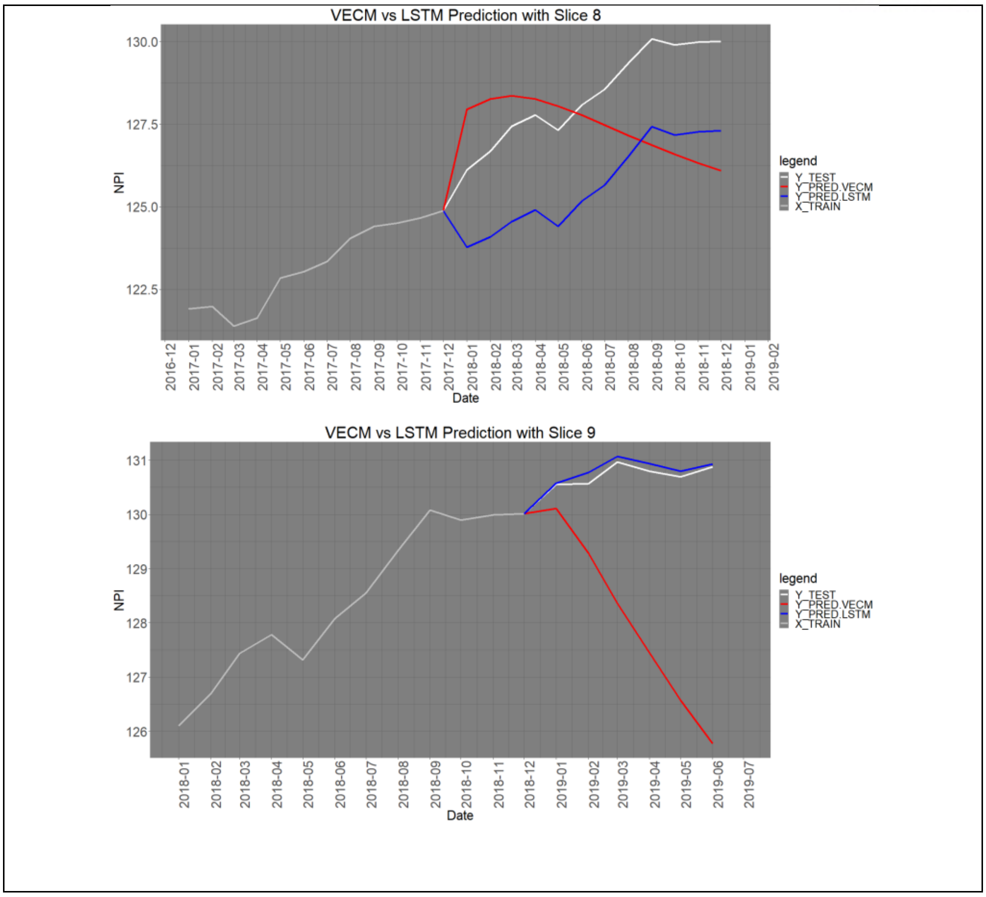


Figure 11. VECM and LSTM prediction comparison from 2010 to 2019 with sliding window test

5. Conclusion

In this article, we focused on the optimal time-series modelling to predict the newbuilding ship price market. We first collected data from January 1986 to June 2019 and verified the causalities between NPI and explanatory variables, including freight rates, interest rates, and second-hand ship prices. Furthermore, model performances were compared to explore the optimal modelling methodology in newbuilding ship price forecasting. The first model, VECM, is used in econometric analysis expended from VAR to address multivariate time-series variables. The second model, LSTM, based on the neural network, which is relatively free from statistical assumptions with advances in machine learning. Further, we proposed a framework that can improve forecasting performance by removing noise using Savitzky–Golay filter in the newbuilding ship price forecasting field.

The variables covered in this study were identified as time-series variables, which were *I(1)* through a unit-root test. We also revealed that economic time-series variables, including freight rates, interest rates, and second-hand ship prices have significant causalities toward newbuilding ship prices.

Next, we proposed an empirically based optimal forecasting methodology in newbuilding ship prices between VECM and LSTM, which represent econometric with linearity assumption and neural network time-series forecasting, respectively. Using RMSE and MAPE as performance indicators, we verified the sliding window test by moving the periods of training data and test data to increase the robustness of predictions. To this end, we first experimented with the cointegration test between the variables and confirmed the existence of long-term equilibrium. Afterward, we performed the VECM model to predict NPI. We also searched for the optimal hyper-parameters of the LSTM model, with the LSTM showing an improved result for RMSE (1.4154), which is 59% lower than VECM (3.472). In particular, LSTM shows better performance than VECM in volatile situations mainly because VECM is an econometric model, which assumes linearity; in contrast, LSTM, being relatively advantageous for non-linear modelling, can be used to make more reasonable predictions in volatile markets.

We also discovered that the Savitzky–Golay filter, which removes noise and minimizes the corruption of time-series data, improves the forecasting performance in the newbuilding ship price market. We explored suitable polynomial orders and window sizes for VECM and LSTM and passed time-series variables into the denoising filter. Next, we found the superiority of the prediction performance that RMSE shows 38% reduction in VECM and 59% in LSTM. These findings suggest LSTM's superiority over VECM and performance improvement when the Savitzky–Golay filter is applied to original data through the empirical analysis of NPI forecasting.

Future Works

Instead of using established theories or objective market analyses, many companies in the shipbuilding industry have drawn on current economic conditions and their own experience to estimate ship prices and make decisions. This attitude poses a danger to the provision of resources, such as labor force and facilities; this attitude also undermines the success of negotiations with shipowners who are willing to purchase ships. Therefore, it is necessary to devise long-term plans for ship markets.

In order to establish scientific forecasting methodologies in practice, more in-depth studies of explanatory variables affecting newbuilding ship prices are needed. In this work, we compared two-time-series models (i.e., VECM and LSTM). The experiment results show that LSTM performs better than VECM. The results, however, preclude generalizing the notion that neural network models are superior to econometric models for forecasting newbuilding ship prices mainly because non-linear econometric time-series models have not been tested to measure performance in newbuilding ship price markets. Therefore, non-linear methodologies such as Tong's (1980) Threshold Auto Regressive (TAR), Engle's (1982) Autoregressive Conditional Heteroscedasticity (ARCH), and Bollerslev's (1986) GARCH should be studied more deeply.

In addition, LSTM, which shows the best performance in this work, is a short-term forecasting model. Hence, follow-up studies such as LSTM sequence-to-sequence for long-term forecasting are necessary (Lee and Jeon (2018)). We hope that future studies will use and benefit from the newbuilding ship price forecasting methodology proposed in this article, helping to lay the foundation for scientific decision-making through forecasting models in the shipbuilding industry.

References

1. Ahn Y., Lee M. (2018). Factor analysis affecting on the charterage of capesize bulk carriers. Korea Trade Review, Vol. 43, No. 3, pp. 125-145.

2. Bae, S., Yu, J. (2018). Predicting the real estate price index using machine learning methods and time-series analysis model. Housing Studies Review, 26(1): 107-133.

3. Beenstock M. (1985). A theory of ship prices. Maritime Policy and Management, Vol. 12, No. 3, pp. 215-225.

4. Beenstock M., Vergottis A. (1989a). An econometric model of the world tanker market. Journal of Transport Economics and Policy, Vol. 23, pp. 263–280.

5. Beenstock M., Vergottis A. (1989b). An econometric model of the world shipping market for dry cargo, freight and shipping. Applied Economics, Vol. 21, pp. 339–356.

6. Beenstock M., Vergottis A. (1992). The interdependence between the dry cargo and tanker markets. Logistics and Transportation Review, Vol. 29, No. 1, pp. 3–38.

7. Beenstock M., Vergottis A. (1993). Econometric modeling of world shipping. London: Chapman & Hall.

8. Bi J., Li S., Yuan H., Zhao Z., Liu H. (2019). Deep neural networks for predicting task time-series in cloud computing systems. Proceedings of the 2019 IEEE 16th International Conference on Networking, Sensing and Control (ICNSC).

9. Bollerslev T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. Journal of Econometrics, Vol. 31, pp. 307-327.

10. Choi Y., Kim H. (2016). The spillover from asset determinants to ship price. Journal of Korea Port Economic Association, Vol. 32, No. 2, pp. 59-71.

11. Engle R. F. (1982). Autoregressive Conditional Heteroskedasticity With Estimates of the Variance of United Kingdom Inflation. Econometrica, Vol. 50, pp. 987-1008.

12. Engle R. F., Granger C. W. J. (1987). Cointegration and error correction representation, estimation, and testing. Econometrica, Vol. 55, pp. 251-276.

13. Geron A. (2017). Hands-on Machine Learning with Scikit-Learn and TensorFlow. O'Reilly Media.

14. Glen D. R. (1997). The market for second-hand ships: Further results on efficiency using cointegration analysis. Maritime Policy and Management, Vol. 24, No. 3, pp. 245-260.

15. Goulielmos A. M. (2019). A brief history of maritime econometrics, 1934-2012. Modern Economy, Vol. 10, pp. 730-756.

16. Granger C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. Econometrica, Vol. 37, 424–438.

17. Granger C. W. J., Newbold P. (1974). Spurious regressions in econometrics. Journal of Econometrics Vol. 2, pp. 111-120.

18. Hale C., Vanags A. (1992). The market for second-hand ships: Some results on efficiency using cointegration. Maritime Policy and Management, Vol. 19, No. 1, pp. 31-39.

19. Haralambides H.E., Tsolakis S. D., Cridland C. (2005). Econometric modeling of newbuilding and secondhand ship prices. Research in Transportation Economics, Vol 12, pp. 65–105.

20. Hochreiter S., Schmidhuber, J. (1997). Long short-term memory. Neural Computation, Vol. 9, No. 8, pp. 1735-1780.

21. Jain, K., Payal. (2011). A review study on urban planning & artificial intelligence. International Journal of Soft Computing and Engineering (IJSCE).

22. Jo, K., Jeong, S., Kim, K., Oh, K. (2018). Scoring model to determine trade timing based on genetic algorithm. Journal of Korean Data & Information Science Society, Vol. 29, pp. 735-745.

23. Johansen S. (1991). Estimating and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. Econometrica, Vol. 59, pp. 1551-1589.

24. Johansen S., Juselius K. (1990). Maximum likelihood estimation and inference on cointegration with applications to the demand for money. Oxford Bulletin of Economics and Statistics, Vol. 52, No. 2, pp. 180-198.

25. Karakitsos E., Varnavides L. (2014). Maritime Economics: A Macroeconomic Approach. Palgrave MacMillan.

26. Kavussanos M. G. and Alizadeh A. H. (2002). Efficient pricing of ships in the dry bulk sector of the shipping industry. Maritime Economics and Logistics, Vol. 29, No. 3, pp. 303-330.

27. Kim M., Lee K., Kim J. (2014). Causality test of the relationship between the freight indexes and the ship prices in second-hand bulk market. Journal of Shipping and Logistics, Vol. 30, No. 3, 2014, pp. 637-654.

28. Kim S., Jung H., Lee H., Yeo G. (2013). An analysis on weighing the decision making factors of ship investments for Korean shipping companies. Journal of Korea Port Economic Association, Vol. 29, No. 2, 2013, 137-157.

29. Kim, E., Oh, K. (2019). Asset allocation strategy using hidden Markov model and genetic algorithm. Journal of Korean Data & Information Science Society, Vol. 30, pp. 33-44.

30. Lee N., Oh K. (2019). KOSPI200 futures index prediction using denoising filter and LSTM. Journal of the Korean Data & Information Science Society 2019, Vol. 30, No. 3, pp. 645-654.

31. Lee, T., Jeon, M. (2018). Prediction of Seoul house price index using deep learning algorithms with multivariate time-series data. SH Urban Research & Insight, Vol. 8, No. 2, pp. 39~56.

32. Lenk M. M., Worzala E. M, Silva A. (1997). High-tech valuation : should artificial neural networks bypass the human valuer?. Journal of Property Valuation & Investment, Vol. 15, No. 1, pp. 8-26.

33. Li J., Parsons M. G. (1997). Forecasting tanker freight rate using neural networks. Maritime Policy and Management, Vol. 24, No. 1, pp. 9-30.

34. Lyridis D. V., Zacharioudakis P., Mitrou P., Mylonas A. (2004). Maritime Economics and Logistics, Vol. 6, pp. 93-108.

35. Park J. (1998). A study on the ordering point of newbuilding by analysis of vessel investment factor. Graduate School of Transportation, Inha University.

36. Pascanu R., Mikolov T., Bengio Y. (2013). On the difficulty of training recurrent neural networks. Proceedings of the 30th International Conference on Machine Learning Vol. 28, pp. 1310-1318.

37. Savitzky A., Golay M. J. E. (1964). Smoothing and differenciation of data by simplified least squares procedures. Analytical Chemistry, Vol. 36, pp. 1627-1639.

38. Schafer R. W. (2011). What is a savitzky-golay filter. IEEE Signal Processing Magazine, Vol. 28, pp 111-117.

39. Selvin S., Vinayakumar R., Gopalakrishnan E. A., Menon V. K., and Soman K. P. (2017). Stock price prediction using LSTM, RNN and CNN-sliding window model. Proceedings of IEEE Int. Conf. Adv. Comput., Commun. Inform (ICACCI), pp. 1643–1647.

40. Stopford, M. (2009). Maritime Economics. 3rd Edition, Rutledge, London.

41. Strandenes S. P. (1984). Price Determination in the time charter and second hand markets. Discussion paper 0584, Norwegian School of Economics and Business Administration, Bergen, Norway.

42. Tong H., Lim K. (1980). Threshold autoregressive, limit cycles and cyclical data. Journal of the Royal Statistical Society Series B, Vol. 42, No. 3, pp. 245-292.

43. Tsolakis S.D., Cridland C., Haralambides H.E. (2003). Econometric modelling of second-hand ship prices. Maritime Economics and Logistics, Vol. 5, pp. 347-377.

44. Veenstra A. W. (1999). Quantitative analysis of shipping markets. T99/3, TRAIL Thesis Series, the Netherlands: Delft University Press.

Acknowledgment

First of all, I would like to thank Heemyung for your active support during my master degree course. Your love and devotion toward our family is the biggest momentum of my life. And my son, Jiwon, I want to be a father who tries to be best in your life. Also, I express my gratitude to my parents, Kyungjin and Heewook, and my sister, Hana. In caring for your son and brother, you have always done your utmost to me. I will emulate your teachings and do my best for my family.

I hope to express my sincere thanks to Professor Hangyun Woo who taught academic and logical thinking, and Professor Daejin Kim and Professor Keeyeun Lee who became committee members. You instructed me what I lacked during my research and gave me directions to be progressed. These lessons will be a guidance for me to become more mature person from now on.

While there are many people that I should acknowledge, forgive me for omissions because of the limitation of space. I hope to give you deep appreciations.